

(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11)特許出願公開番号  
特開2001-143445  
(P2001-143445A)

(43)公開日 平成13年5月25日(2001.5.25)

(51)Int.Cl. <sup>7</sup>	識別記号	F I	テーマコード <sup>*</sup> (参考)
G 1 1 B 27/031		G 1 1 B 27/02	A
H 0 4 N 5/91			Z
		H 0 4 N 5/91	N

審査請求 未請求 請求項の数12 O L (全 12 頁)

(21)出願番号	特願2000-274132(P2000-274132)	(71)出願人	593081408 ソニー・ユナイテッド・キングダム・リミ テッド Sony United Kingdom Limited イギリス国 サリー, ウェブリッジ, プ ルックランズ, ザ ハイツ (番地なし)
(22)出願日	平成12年9月8日(2000.9.8)	(74)代理人	100067736 弁理士 小池 晃 (外2名)
(31)優先権主張番号	9921234.2		
(32)優先日	平成11年9月8日(1999.9.8)		
(33)優先権主張国	イギリス (GB)		

最終頁に続く

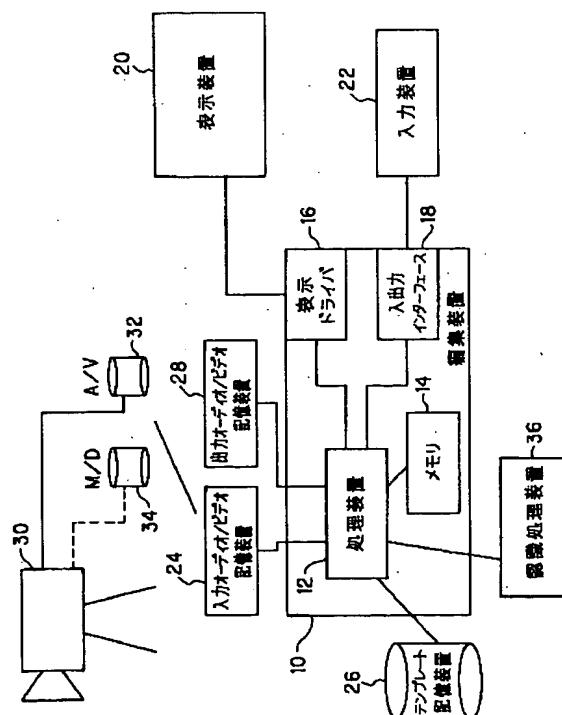
**最終頁に続く**

(54) 【発明の名称】 編集装置及び方法

(57) 【要約】

【課題】 ビデオ及び／又はオーディオ素材のようなソースコンテンツを編集し、編集コンテンツシーケンス、ソースコンテンツの部分に対応するソースメタデータを作成する編集装置を提供する。

【解決手段】 編集装置は、コンテンツの所望のスタイルを規定するテンプレートを格納するテンプレート記憶装置と、テンプレートにおいて所望のスタイルを規定するテンプレートメタデータと、ソースメタデータをテンプレートメタデータと比較することによりテンプレートをソースコンテンツに適用し、この比較に応じてソースコンテンツの部分の配列して編集コンテンツシーケンスを作成する処理装置とを有する。これによって、編集者に第1カットでコンテンツが所望のスタイルに配列される効率的な技術が提供されるので、編集者はソースコンテンツの編集を行うに前にコンテンツに理解するために時間を割く必要がなくなる。ユーザは、第1カットから適切であると思われる変更をさらに進めることができる。



**【特許請求の範囲】**

【請求項1】 ソースコンテンツを編集して編集コンテンツシーケンス、上記ソースコンテンツの部分に対応するソースメタデータを作成する編集装置において、  
所望のスタイルのコンテンツ、テンプレートにおいて所望のスタイルを規定するテンプレートメタデータを格納する記憶装置と、

上記ソースメタデータを上記テンプレートメタデータと比較することによって上記テンプレートを上記ソースコンテンツに適用し、上記比較に応じて上記ソースコンテンツの部分に配列して編集コンテンツシーケンスを作成する処理装置とを備える編集装置。

【請求項2】 上記ソースコンテンツの部分に配列して選択からなる編集コンテンツシーケンスを生成し、上記ソースメタデータと上記テンプレートメタデータとの比較に基づいて上記ソースコンテンツの部分のサブセットを順序付けることを特徴とする請求項1記載の編集装置。

【請求項3】 上記テンプレートは、上記所望のスタイル、対応するテンプレートメタデータを有する各離散セクションを規定するように連係された複数の離散セクションを規定することを特徴とする請求項1記載の編集装置。

【請求項4】 上記ソースコンテンツは、ビデオ素材、ビデオ素材の種類を特定する多数の上記離散セクション、少なくともビデオ素材に対応する種類を特定するメタデータからなる離散セクションに対応する上記テンプレートメタデータとからなることを特徴とする請求項3記載の編集装置。

【請求項5】 上記コンテンツは、オーディオ素材、上記オーディオ素材に関連するメタデータを含む多数の上記離散セクションに対応する上記テンプレートメタデータからなることを特徴とする請求項3記載の編集装置。

【請求項6】 上記ソースメタデータを上記テンプレートメタデータと比較することにより、上記テンプレートを上記ソースコンテンツに適用するときに生成されて上記処理装置によって用いられる上記ソースコンテンツ、上記ソースメタデータに対応するソースメタデータを生成するために、1以上の所定の認識ツールにアクセスし、上記ソースコンテンツを上記1以上の所定の認識ツールによってレビューさせるメタデータ処理要素を備えることを特徴とする請求項1記載の編集装置。

【請求項7】 上記メタデータの処理要素は、ユーザ入力に応じて上記ソースメタデータの生成の間にユーザとの対話を可能にすることを特徴とする請求項6記載の編集装置。

【請求項8】 上記テンプレートは、コンテンツ、上記テンプレートにおいて上記所望のスタイルを規定するテンプレートメタデータを規定することを特徴とする請求

項1記載の編集装置。

【請求項9】 上記所望のスタイルを規定するために連係された複数の離散セクション、対応するテンプレートメタデータを有する各離散セクションからなることを特徴とする請求項8記載の編集装置。

【請求項10】 編集装置を操作してソースコンテンツを編集し、編集コンテンツを作成する編集方法であって、

(a) 編集されるソースコンテンツ、上記ソースコンテンツの部分に対応するソースメタデータを受けとるステップと、

(b) コンテンツ、上記テンプレートにおいて所望のスタイルを規定するテンプレートメタデータを規定するテンプレートを提供するステップと、

(c) 上記ソースメタデータを上記テンプレートメタデータと比較することで上記テンプレートを上記ソースコンテンツに適用するステップと、

(d) ステップ(c)での比較に応じて上記ソースコンテンツの部分に配列し、上記編集コンテンツシーケンスを生成するステップとを有する編集方法。

【請求項11】 ソースコンテンツを編集して編集コンテンツシーケンス、上記ソースコンテンツの部分に対応するソースメタデータを編集する編集装置を操作するコンピュータプログラムであって、上記コンピュータプログラムは上記編集装置に、

(a) 記憶装置からコンテンツ、テンプレートにおいて所望のスタイルを規定するテンプレートメタデータを検索するステップと、

(b) 上記ソースメタデータを上記テンプレートメタデータと比較することで上記テンプレートを上記ソースコンテンツに適用するステップと、

(c) ステップ(b)での比較に応じてソースコンテンツの部分に配列して上記編集コンテンツシーケンスを作成するステップとからなる操作を実行させることを特徴とするコンピュータプログラム。

【請求項12】 ソースコンテンツを編集して編集コンテンツシーケンス、上記ソースコンテンツの部分に対応するソースメタデータを編集する編集装置で用いられるテンプレートにおいて、

所望のスタイルを規定するように連係する複数の離散的なセクションからなり、上記離散セクションは対応するテンプレートメタデータを有することを特徴とするテンプレート。

**【発明の詳細な説明】**

【0001】

【発明の属する技術分野】 本発明は、ソースコンテンツ(source content)を編集して編集コンテンツシーケンス(edited content sequence)を作成する編集装置及び方法、コンピュータプログラム並びにテンプレートに関する。ソースコンテンツは、例えばビデオ及び/又は

オーディオ素材 (video/audio material) である。

【0002】

【従来の技術】ソースコンテンツの編集は、従来から負担が大きい作業であり、熟練した編集者によって編集装置 (editorial suite) で行われる。編集装置 (editing system) は、通常は編集者の管理下でソースコンテンツを編集するのに用いられる。現在の編集装置は洗練されたツールを有し、編集者は、ソースコンテンツを複雑な編集効果で処理することができる。

【0003】

【発明が解決しようとする課題】しかし、編集者は、このような編集の前に、ソースコンテンツをレビュー (review) して編集する素材を理解するために相当の時間を要している。編集作業の効率は、最初にソースコンテンツを理解するために時間を費やすことで、悪くなっている。

【0004】本発明は、上述の実情に鑑みて提案されるものであって、編集の前に編集者が素材を理解するために費やす時間を削減することにより、編集効率を向上させるような編集装置及び方法、コンピュータプログラム並びにテンプレートを提供することを目的とする。

【0005】

【課題を解決するための手段】本発明は、ソースコンテンツを編集して編集コンテンツシーケンス、ソースコンテンツの部分 (portion) に対応するソースメタデータ (source metadata) を作成する編集装置を提供する。この編集装置は、所望のスタイルのコンテンツ、所望のスタイルを規定するテンプレート (template) におけるテンプレートメタデータ (template metadata) を格納する記憶装置 (storage) と、ソースメタデータ (source metadata) をテンプレートメタデータと比較することでテンプレートをソースコンテンツに適用し、この比較に応じてソースコンテンツの部分を実列して編集コンテンツシーケンスを作成する処理装置 (processing unit) とを備える。

【0006】本発明は、コンテンツの所望のスタイルを規定するテンプレート、テンプレートにおいて所望のスタイルを規定するテンプレートメタデータを提供する。本具体例において、テンプレートは、所望のスタイルを規定するように連係 (link) する複数の離散的なセクションからなり、各離散セクションは対応するテンプレートメタデータを有している。従って、本発明のこの観点によると、多数のテンプレートが用意されて編集コンテンツシーケンスに必要な特定のフォーマットを規定し、これらのテンプレートは本発明の編集装置の使用に適するように選択される。

【0007】本発明は、ソースコンテンツを編集して編集ソースコンテンツを作成するように編集装置を操作する編集方法を提供し、この編集方法は、(a) 編集されるソースコンテンツ、ソースコンテンツの部分に対応す

るソースメタデータを受け取るステップと、(b) コンテンツ、テンプレートにおいて所望のスタイルを規定するテンプレートメタデータを提供するステップと、

(c) ソースメタデータをテンプレートメタデータと比較することでテンプレートをソースコンテンツに適用するステップと、(d) ステップ (c) の比較に応じてソースコンテンツの部分を実列して編集コンテンツシーケンスを作成するステップとを有する。

【0008】本発明は、ソースコンテンツ、ソースコンテンツに部分に対応するソースコンテンツメタデータを編集して編集コンテンツシーケンスを作成するように編集装置を操作するコンピュータプログラムを提供し、コンピュータプログラムは編集装置に、(a) 記憶装置から所望のスタイルを規定するテンプレート、テンプレートにおける所望のスタイルを規定するテンプレートメタデータを検索するステップと、(b) ソースメタデータをテンプレートメタデータと比較することでテンプレートをソースコンテンツに適用するステップと、(c) ステップ (b) における比較に応じてソースコンテンツの部分を実列して編集コンテンツシーケンスを作成するステップとを実行させる。

【0009】

【発明の実施の形態】以下、本発明に係る編集装置及び方法、コンピュータプログラム並びにテンプレートの実施の形態について、図面を参照しながら詳細に説明する。

【0010】図1は、本発明を適用した編集装置10の具体的な構成を示すブロック図である。この図1に示すように、編集装置10は、編集コンテンツを作成するためにソースコンテンツを受け取って編集操作をソースコンテンツに施す処理装置 (processing unit) 12、例えばマイクロプロセッサを備えている。内蔵メモリ14は、例えばリードオンリーメモリ (read only memory:

ROM)、ランダムアクセスメモリ (random access memory: RAM) などであり、処理装置12で使用されるデータとともに、処理装置12で実行されるソフトウェアを格納する。

【0011】編集装置10は、表示ドライバ16を備え、表示ドライバ16は、表示信号を編集装置10に接続された表示装置20に出力する。また、編集装置10は、入出力インターフェース18を備え、入出力インターフェース18は、入力装置22、例えばキーボード、マウス、ジョグシャトルダイヤル、ペンとタブレットなどを編集装置10に接続するインターフェースを司る。編集装置10のユーザは、表示装置20及び入力装置22を介して、編集装置10によって実行される編集処理と対話 (interact) することができる。

【0012】本具体例では、編集装置10によって編集されるソースコンテンツは、デジタルカメラ30から出力されたようなオーディオ/ビデオ (audio/video) 以

下、A/Vという。)素材であるとされる。A/V素材は、アナログデータでも作成されるが、本具体例の編集装置10はデジタルデータによって動作するので、アナログA/V素材は編集装置10への入力される前にデジタルデータに変換される。

【0013】A/V素材がデジタルカメラ30によって生成されると、適切な媒体32、例えばデジタルテープ又はデジタルディスクに格納される。最近のデジタル技術では、A/V素材が最初に作成された時点において、あるメタデータがA/V素材に対応させられる。本具体例では、このようなメタデータは、適切な媒体32として、例えばデジタルテープ又はディスクに格納される。なお、メタデータは、対応するA/V素材と同一の物理記憶媒体32に記録されるが、別の物理記録媒体34にも記録することもできる。

【0014】A/V素材が編集装置10に送られる前のある時点において、例えば手作業でメタデータを加えることによって、追加のメタデータを加えることもできる。A/V素材が編集される時、A/V素材は、対応するメタデータ(以下では、ソースメタデータと呼ぶ。)とともに、入力A/V記憶装置(input A/V store)24、例えばデジタルテーププレーヤ、デジタルディスクプレーヤ又はサーバに格納される。

【0015】本具体例では、1以上のテンプレートが記憶装置、例えばデータベースやネットワークを介してアクセスされる記憶装置から提供される。ここで、図1ではテンプレート記憶装置26は別個の構成要素として表されているが、テンプレートは編集装置10において局所的に、例えばメモリ55に格納することもできる。多くの種類の編集素材、例えばニュース報告は、似通ったパターンに従うことが認められる。したがって、本具体例では、各テンプレートはA/V素材の所望のスタイルを規定し、テンプレートメタデータは各テンプレートにおいて所望のスタイルを規定する。

【0016】編集装置10は、テンプレート記憶装置26から適切なテンプレートを検索し、テンプレートメタデータをソースA/V素材に対応するソースメタデータと比較することで、テンプレートを入力A/V記憶装置24から供給されるソースA/V素材に適用するようになされている。この作業は、処理装置12によって実行され、ソースA/V素材は、ソースメタデータとテンプレートメタデータの比較に応じて、編集コンテンツシーケンスを作成するよう配列される。編集コンテンツシーケンスは、処理装置12から出力A/V記憶装置28へ出力される。この出力は、編集装置10のユーザによってさらに変更することができる。入力A/V記憶装置24と同様に、出力A/V記憶装置28は、デジタルテープ、デジタルディスクプレーヤ又はサーバなどから構成される。

【0017】本具体例において、テンプレートは、編集

するA/V素材に対するユーザの理解に基づいて、編集装置のユーザにより選ばれると想定される。例えば、ユーザは、A/V素材の編集を求めている人からテンプレートの選択について明示的な指示を与えられる。したがって、例えば、編集装置10のユーザは、テンプレート記憶装置26に格納された特定のニュース向けのテンプレートを使用するという指示とともに、ニュース事項に関係するあるA/V素材を受け取る。なお、ユーザは、編集するA/V素材のタイプに基づいて、使用するテンプレートを自分で決めることもできる。テンプレートは、テンプレート記憶装置26から具体的には手動で選択されるが、編集装置10によるテンプレートの自動選択のように、編集するA/V素材に対応するメタデータが、どのテンプレートを使用するか処理装置12に直接に指示することも考えられる。

【0018】本具体例では、編集装置10は、1以上のプラグイン認識ツールの形態を有する多くの認識処理装置36に接続することができる。これらの認識ツールは、編集装置10によってメタデータの生成の補助に用いられる。例えば、顔を検出する認識ツールは、ソースA/Vをスキャンしてビデオに含まれる顔を探すために用いられる。一旦これらの顔が識別されると、これらは編集装置10のユーザに提示され、ユーザが検出された顔の識別に関係するメタデータを入力するように促す。なお、認識ツール又は編集装置10の何れかが人の顔の適切なデータベースにアクセスできるなら、顔の自動的な認識もできる。同様に、音声を認識し、A/Vソースにおけるオーディオ素材に関連するメタデータの追加ができるような認識ツールも提供される。

【0019】次に、編集装置10によって用いられるテンプレートを、図2を用いて説明する。図2は、テンプレート記憶装置26に格納された具体的なテンプレートを示す図である。A/Vシーケンスを記述するテンプレートは、5つ主要セクション40、42、44、46、48からなる。第1セクション40は導入(introduction)を提供する素材からなり、第2セクション42は1以上の静止画(still)からなり、第3セクション44は特定のショット(shot)の場面を与え、第4セクション46はさらにショットの場面を与え、最後のセクション48は終了シーケンスを提供する。さらにセクション50のボイスオーバー(voice over)は、素材にオーディオボイス(audio voice)を与え、セクション40、44及び46に対応するオーディオを与えるために用いられる。

【0020】(ビデオワイプ(video wipe)、ページターン(page turn)、オーディオフィード(audio fade)、クロスフェード(cross fade)などのような)多数の特定の編集効果、キーイング及び/又は技術的な調整は、各セクションに対応し、編集コンテンツシーケンスの生成の間、ユーザは、これら対応する効果、キー

ングまたは調整のどれを使うか入力を催促される。例えばセクション40は、(ブロック41に示すように)導入の最後に適用される右上から左下へのボックスワイプ(box wipe)を示す編集効果に対応し、セクション42は、(ブロック43に示すように)静止画の最後に適用される左下から右上へのボックスワイプを示す編集効果に対応している。同様に、ストレートカット(straight cut)は、ブロック45及び47の間に示されるように、セクション44及び46の間、及びセクション46及び48の間で特定される。また、ブロック52に示されるように、オーディオフィードアップ(audio fade up)はセクション50の最初に特定され、ブロック52及び53に示すように、ビデオフェードイン(video fade in)及びビデオフェードアウト(video fade out)

はセクション46のキャプションに対応している。例えばセクション40及び46にそれぞれ対応するブロック54及び55に示すように、ユーザにキャプションが提案される。

【0021】各セクションは、対応する1以上のメタデータの事項を有する。このようなメタデータは、以下でテンプレートメタデータと呼ばれる。特定のセクションに対応するテンプレートメタデータは、そのセクションに対応するコンテンツの種類を規定するのに用いられる。メタデータには、テンプレートの各セクションに対応する多くの種類がある。次の表1は、メタデータの種類のある例を示す。

【0022】

【表1】

ビジネス情報：カメラマン、ジャーナリスト、割当番号、版權	トランジションマーク：カット、ディゾルブ、フェードの自動識別及びカテゴリ化
装備仕様：カメラシリアル番号、レンズパラメータ、技術特徴	ショット構成マーク：ズーム及び広角度マーク
素材識別：固有素材識別子(UMID)	雑音マーク：高周雑音(おそらく高電力低周波数雑音)又は低周雑音(すなわち無感部分)
位置マーカ：GPS、時間帯、固定位置(例えば、イギリス国会の議場のマイクロフォン、ゴール端のカメラ)	テンポ/ビートマーク：リズム、周期的音響イベント
ユーザ定義マーカ：グッドショットマーカ、制作コメント	色/コントラスト変化：カメラの閃光、電光効果によって引き起こされる。
外部運動：ズーム又はパニングによる大域的な運動	ピッチ(音の高さ)/ボリューム(音の太さ)変調：これは、音楽や対話のような長期間のオーディオイベントに当てはまる。
内部運動：スポーツにおける、又は人の動き(すなわち、「比較的」遅く変化する運動を伴う人のような凝集した被写体)のような被写体の軌跡、	オーディオパターン/不連続：リズム、テンポ、歩調のようなもの。

【0023】最近、メタデータ、コンテンツを記述するメタデータの生成に関心が持たれている。メタデータとコンテンツの対応関係は、保管(archive)の際に有益であると考えられている。メタデータは、所望のコンテンツの後の探索(search)及び検索(retrieve)の処理における補助に用いられる。メタデータには、例えば、SMPTE-EBU(Society of Motion Picture and Television Engineers - European Broadcast Union)、MPEG-7(Motion Pictures Expert Group)のように、多くの規格が開発されている。ここで、Motion Pictures Expert Groupは、ISO/IEC standards body SG29\WG11を意味する。

【0024】一定量のメタデータ、例えば(良いショットであることを示す)グッドショットマーカ(good shot marker)、(どこで記録を開始/停止したかを示す)録音マーカ、GPS(global positioning system)位

置、日時、固有素材識別子(unique material identifier: 以下、UMIDという。)、カメラ設定、焦点、ズームなどは、ソースに自動的に追加することができる。これに加えて、コンテンツからメタデータを自動抽出する技術の進歩は大きく、例えば顔及び特徴検出ツール、音声認識ツールなどの認識ツールなどを利用することで、コンテンツの特徴を識別し、適切なメタデータを追加することができる。

【0025】本発明は、ソースコンテンツに対応するメタデータ(以下では、ソースメタデータという。)技術の最近の発展を利用し、編集処理の効率を改善することを図っている。具体的に、本発明では、テンプレートは、コンテンツ、テンプレートにおいて所望のスタイルを規定するテンプレートメタデータの所望のスタイルを規定する。テンプレートメタデータは、本質的にソースメタデータの選択又はソースメタデータの適合基準を表

す。ハードディスク、フロッピー（登録商標）ディスク、CDROMなどに格納されたテンプレートは、編集装置によって検索され、ソースメタデータをテンプレートメタデータと比較することでテンプレートをソースコンテンツに適用するようになっている。ソースコンテンツの部分は、この比較に応じて配列され、編集コンテンツシーケンスが作成される。

【0026】本発明においては、多くの種類の編集素材、例えばニュース報道、スポーツクリップ／概要、ビデオモンタージュ、ドキュメンタリなどは、かなり類似のパターンに従うものとされる。例えば、国会における討論の報道は、導入（introduction）に始まり、国会議事堂の外側のショット（shot）、政治家の討論のカット（cut）、再び国会議事堂の他のカットに進み、最後に終わりのコメントを述べるレポートに戻る。ショットに対して適用される編集効果、又はショットの間で適用される編集効果も、類似のパターンに従っている。本発明は、このような事実と、メタデータが入手しやすくなっていることを利用し、様々な所望のスタイルに対して作成されたテンプレートについて、これらのテンプレートにおけるテンプレートメタデータが所望のスタイルを規定する編集装置を提供する。所望のスタイルは、ショットの一般的なシーケンス、ソースコンテンツから部分を選択するのに用いる基準、適用される編集効果などのような特徴を特定する。

【0027】編集コンテンツシーケンスを作成するために、ソースメタデータをテンプレートメタデータと比較することで、選ばれたテンプレートがソースコンテンツに適用される。この編集コンテンツシーケンスは、編集者にソースコンテンツが所望のスタイルで配列された「第1カット（first cut）」を提供するので、ユーザがソースコンテンツの編集を行う前にコンテンツを理解するために時間を割く必要がなくなる。この第1カットから、ユーザは適当であると思われる変更に進むことができる。

【0028】ソースコンテンツを配列して編集コンテンツシーケンスを作成するには、1以上のステップを要する。例えば、ソースコンテンツの配列は、ソースコンテンツの部分の記録、ソースコンテンツからの部分の選択、ソースコンテンツへの編集効果の適用、キャプションの挿入などのステップを1以上含む。本具体例では、ソースコンテンツの部分の配列する編集コンテンツシーケンスの作成は、少なくとも、ソースメタデータとテンプレートメタデータとの比較に基づくソースコンテンツの部分のサブセットの選択または順序付けを有する。したがって、例えば、本具体例では、ソースコンテンツは同じショットの多重「テーク（take）」を含み、ソースメタデータはどのテークが好ましいかを例えばテークに対応するグッドショットマーカを用いて示す。グッドショットマーカがテンプレートメタデータによって特定さ

れると、テンプレートメタデータとソースメタデータとの比較は、編集コンテンツシーケンスの作成において、特定のショットから使用する所望のテークだけを選択するために用いられる。このアプローチにより、原ソースコンテンツ（original source content）よりかなり短い編集コンテンツシーケンスが作成されると考えられる。

【0029】テンプレートは様々な形を取るが、本具体例では、テンプレートは所望のスタイルを規定するために連携する複数の離散セクション（discrete section）を規定し、各離散セクションは対応するテンプレートメタデータを有している。各離散セクションは、本具体例では、所望のシーケンスの部分に対応する所望のコンテンツのタイプを規定すると考えられ、テンプレートメタデータの1以上の項目（item）は、各離散セクションに対応してこの離散セクションに対応するコンテンツの種類を規定する。

【0030】本具体例において、ソースコンテンツは、ビデオ素材からなり、多数の離散セクションはビデオ素材の種類を特定し、このような離散セクションに対応するテンプレートメタデータは、少なくとも、対応するビデオ素材の種類を特定するメタデータからなる。たとえば、1つの離散セクションは、「静止（still）」ショットとされ、この離散セクションに対するメタデータは、たとえばソースコンテンツにおける静止ショットを特定する種類である。このアプローチによって、この離散セクションのテンプレートメタデータがソースメタデータと比較されると、ソースコンテンツにおける静止ショットが特定され、これらの静止ショットを表すソースコンテンツの1以上の部分は、編集コンテンツシーケンスの作成に用いられる離散セクションに対応させられる。

【0031】例えば、ビデオ素材に加えて、又はビデオ素材とは別に、コンテンツはオーディオ素材からなり、多数の離散セクションに対応するテンプレートメタデータはオーディオ素材に関連するメタデータを含む。このようなメタデータは、例えば高い暗騒音（background noise）、低い暗騒音、リズム、周期的音響イベントなどを表し、テンプレートメタデータをソースメタデータと比較して所望のタイプのオーディオを含むソースコンテンツの部分の特定するときに使うことができる。

【0032】上述したように、メタデータは、ソースコンテンツにソースにおいて加えられるか、ソースコンテンツが記録された後のある段階で例えば手動で加えられる。これに加えて、本具体例では、編集装置は、ソースメタデータを生成してソースコンテンツに対応させるために1以上の所定の認識ツールにアクセスし、ソースコンテンツを1以上の認識ツールによってレビューさせるメタデータ処理要素と、このように生成され、ソースメタデータをテンプレートメタデータと比較することによ

ってテンプレートをソースコンテンツに適用するとき、処理装置によって使用されるソースメタデータとから構成される。メタデータ処理要素(metadata processing element)は、ハードウェア、ファームウェア及び/又はソフトウェアに、1以上の所定の認識ツール(recognition tool)へのアクセス、及びこれらのツールを用いたメタデータの生成の制御に適するように組み込まれる。したがって、例えば顔を識別する認識ツールは、ビデオをスキャンしてビデオに含まれる顔を探すように使用される。一旦これらの顔が識別されると、自動又は手動による顔の認識が用いられて関係ある人を特定し、認識に応じた適切なメタデータを加える。

【0033】本具体例において、メタデータ処理要素は、ユーザ入力に応じてソースメタデータの生成の間にユーザとの対話(interaction)をすることができる。したがって、上記の例では、見いだされた顔がユーザに提示され、ユーザはこれらの画像に関するメタデータ、例えば画像に示された人物を識別するように促される。

【0034】特定のテンプレートが選択されると、編集装置10の処理装置12は、編集コンテンツシーケンスを生成するために、テンプレートをソースA/V素材に適用する。これは、ソースA/V素材の個々の部分をテンプレートの特定のセクションに対応させるため、ソースメタデータをテンプレートメタデータと比較することにより行われる。例えば、ソースA/V素材の特定の部分は、セクション42に対応するテンプレートメタデータに対応するソースメタデータを有し、A/V素材のこれらの部分が静止画に関係していることを特定する。同様に、A/V素材の他の部分は、第1の特定のショットを特定するセクション44に対応するテンプレートメタデータに対応するソースメタデータを有している。また、ソース素材のある部分は、ボイスオーバー素材に関連することを特定するソースメタデータを有し、セクション50に関連することと特定するソースメタデータを有し、メタデータのセクション50に関連するテンプレートメタデータとの対応は、ソース素材をテンプレートのセクション50に対応させるために用いられる。

【0035】テンプレートメタデータとソースメタデータとの比較に応じて、ソースA/V素材の関連する部分は、例えば比較に応じて配列され、テンプレートによって規定される一般スタイルに従う編集コンテンツシーケンスを生成する。これは、好ましくはソース素材の部分のサブセットの選択を含み、これらの部分を再び順序付けてテンプレートによって規定されたスタイルに従うようにする。特定されたどんな編集効果なども、この時点で適用される。したがって、適切なテンプレートを選択することによって編集装置のユーザは自動的に編集シーケンスの第1カットを提供され、ソースコンテンツの所望の部分は選択されて順序付けられ、編集効果が適用される。したがって、ユーザは編集コンテンツシーケ

スをレビューし、さらに変更する必要があるかどうか判定することができる。

【0036】使用に適切な追加の場面(footage)を検索するため、編集装置10がある保管素材(archive material)を探索するように、あるタグを特定のセクションに対応させることもできる。この探索は、このセクションに関連する適切な場面を検索するために、例えばタグが対応する特定のセクションのテンプレートメタデータと保管メタデータに対応するメタデータを比較することで開始される。図2の例を参照すると、この処理はセクション46に対応するブロック46によって示され、注記の形のソースA/V素材に対応するソースメタデータは特定の保管場面の選択を有することを示している。

【0037】テンプレートのソースコンテンツへの適用は、図3から図4の例を用いて説明される。図3は、ニュース事項に関係のある特定のソースA/V素材の制作者により特定されるニューステンプレートAを示している。ニューステンプレートAに応じて、所望のスタイルは、6つのセクション61、62、63、64、65及び66によって規定される。第1セクション61は報告するニュース項目についての導入に関連し、第2セクション62は国会議事堂の外のショットを与え、第3セクション63はレーポータのショットを与え、第4セクション64は政治家のショットを与え、第5セクション65は国会議事堂の静止画を与え、最後のセクション350はこれまでのトピックについて建物の外での終わりのコメントを与える。図2を参照して記述されたテンプレートと同様に、特定の編集効果などは各セクションに対応するが、簡単のためにこれらは図3では省略されている。各セクションは、対応する多数のテンプレートメタデータの項目を有し、ソースコンテンツをレビューしてどのテンプレートのセクションにソースコンテンツの特定の部分が対応させられるかを決定する。

【0038】図4のAは、ソース媒体に現れるように、原ソースA/V素材を示している。この例では、ソース素材は、個々のショット又はキャプチャされたテークに関連する10個の部分からなる。ある部分の間で、カメラ設定の間に記録された素材のように、具体的に興味のない雑多な素材があり、この素材は図4の領域aによって表されている。具体的な先行技術の編集装置において、編集装置のユーザは、(しばしばソースコンテンツの制作に関わった誰か、例えばジャーナリストに伴われ)、編集コンテンツシーケンスの最終版の特定の部分にどの部分が用いられるか決めるために、原ソースを数回スキャンしてソース素材の全コンテンツを理解しなければならない。しかし、本具体例では、ソースコンテンツの制作に関わった誰か、例えばプロデューサは、編集コンテンツシーケンスは、ニューステンプレートAに従うことを特定する。これに応じて、ニューステンプレートAは編集装置10により使用されるテンプレート記憶

装置26から選択される。

【0039】テンプレートをソースコンテンツへの適用する前に、ソースメタデータは、ソースコンテンツの様々な部分に対応させられる。ソースメタデータは、例えば最初にキャプチャされたときにコンテンツに対応させられ、又は適切なメタデータの手動の挿入によって後で加えられる。対応するソースメタデータとともに、ソースコンテンツを編集装置10が受け取ると、編集装置10は、テンプレートAの各セクションに対応するテンプレートメタデータをソース素材の部分に対応するソース素材と比較することで、ニューステンプレートAをソース素材に適用する。

【0040】このように、図3において、導入セクション61は、ソース素材における導入部分を一般に特定するタイプのテンプレートメタデータを含み、ソースメタデータのレビューにおいては、図4に示すように、セクション61に対応するテンプレートメタデータの項目は導入の2つのテークを表す部分74及び75に対応するソースメタデータに対応することが見いだされる。例えば部分74及び75は、これらの部分を導入として特定するソースメタデータの手作業による挿入によって導入することで、テークであると識別される。

【0041】なお、例えば、これらの部分は、例えばグッドショットマーカの使用によって、導入の2つのテークのどちらが好ましいか識別するメタデータを含む。セクション61に対応するテンプレートメタデータがソース素材において良好としてマークされた部分の識別をサポートするテンプレートを含むこともできれば、テンプレートをソースコンテンツに適用することによって導入の所望のテークが選択されることになる。したがって、図4の例において、このようなソースメタデータが加えられ、好ましい第1のテークを表す導入1の部分74が好ましいとされる。同様に、部分71はセクション61に対応している。

【0042】同様に、セクション62に対応するテンプレートデータは、具体的にはメタデータであり、国会議事堂のショットの特定をサポートする。例えば、国会議事堂のGPS位置を識別するメタデータは、セクション62に対応している。ソースコンテンツをレビューするときに、GPSメタデータが部分70及び71に対応していることが見いだされると、これらの部分は、制作される編集コンテンツシーケンスにおける部分の順序を決定する目的で、部分62に対応させられなければならない。なお、本具体例においては、ソースメタデータは、ショットの1つをより良好であるとして特定し、この場合には部分71によって表されるショット、テンプレートメタデータは、ソースメタデータを介して良好なショットを識別する種類を含む。したがって、ショット71が、セクション62に対応する。

【0043】同様に、レポートと政治家にそれぞれ関係

するセクション63及び64に関して、メタデータは、レポートのテークに関係するソースコンテンツの部分と、政治家のテークに関係するコンテンツの部分とを特定するソースコンテンツに加えられる。これは、直接に適用されるが、ユーザには、顔のショットを提供するソース素材のこれらの部分を識別するために使われ、これらの部分を適切なメタデータでラベル付けしてレポートに関係する部分と政治家に関係する部分とを特定する顔検出ツールのような認識ツールが提供されている。したがって、ソース素材をレビューするとき、編集装置10は、部分76及び77がセクション63に適切に対応し、部分72及び73がセクション64に適切に対応することを特定する。しかし、この例では、グッドショットマーカは、部分76に対応するので、部分77より優先して選択される。

【0044】この例では、グッドショットマーカは部分72及び73に対応することではなく、好ましい具体例では両方の位置はセクション64に対応し、ユーザは後で2つのテークの選択ができる。これは、例えばテンプレートメタデータがグッドショットマーカなどの特定をサポートする種類を含まない場合のように、セクション64が好ましいショット又はテークの識別ができるテンプレートメタデータのタイプを含まない場合に当てはまる。

【0045】適切なメタデータは、セクション65及び66に対応し、静止画像、又は外でのコメントのそれぞれに関係するソースコンテンツのどの部分の識別もサポートする。したがって、テンプレートメタデータをソースメタデータと比較することにより、図3に示されたニューステンプレートAが図4のAに示されたソースコンテンツに適用され、再配列されたソースコンテンツが図4のBで特定される編集ソースコンテンツが得られる。

【0046】ソースコンテンツをニューステンプレートAと比較することによって編集装置10によって自動的に生成された編集コンテンツシーケンスは、編集装置のユーザが必要な調整をさらに行えるように表示される。本具体例では、ソースコンテンツのある部分が好ましい二者択一のテーク/ショットで優先して落とされ、断片化された様々なソースコンテンツの部分が最終編集素材の適切な順序に再配列されるので、ユーザがソースコンテンツの編集を行う前にコンテンツを理解するために時間を割く必要がなくなると考えられる。この第1カットから、ユーザは、適切であると思われるさらなる変更の作成にさらに進むことができる。これによって、編集処理の効率は著しく改善される。

【0047】図5は、第1の具体例において、編集装置10で行われる処理を示すフローチャートである。ステップ81において、ソース素材が供給され、例えば入力A/V記憶装置24に読み込まれる。ステップ82において、編集装置10は、テンプレートが選択されている



かどうかを判定する。本具体例では、ユーザが入力装置22を介して適切な選択をすることによって手動で行われる。なお、上述したように、使用するテンプレートを特定し、編集装置10によって直接に認識できるメタデータは、ソース素材にすでに対応しているので、編集装置10が自動的に所望のテンプレートを選択することができる。

【0048】処理はステップ83に進み、編集装置10は、テンプレートをテンプレート記憶装置26から検索し、テンプレートメタデータをソースメタデータと比較することによりソース素材に適用する。比較の結果、ステップ84において、ユーザによるレビューのために編集コンテンツシーケンスが作成される。上述したように、編集コンテンツシーケンスは、ステップ83で実行したテンプレートメタデータとソースメタデータの比較に応じて、原ソース素材の部分を配列することにより作成される。例えば、この処理は、ソース素材の部分の選択及び順序付けを含む編集コンテンツシーケンスを生成するが、編集効果のアプリケーションなども含む。

【0049】図5では、ソース素材が適切なメタデータを含むテンプレートのソース素材への適用するものとしている。図6は、ソースメタデータが編集処理の間に補給される第2の具体例を示す。

【0050】図6の最初の3つのステップ81、82及び83は、前に図5で示したものと同様である。しかし、テンプレートメタデータをソースメタデータと比較することにより、一旦テンプレートがソース素材に適用されると、ステップ91においてテンプレートの部分が占められていない(not populated)こと、すなわちどのセクションもこれらに対応するソース素材の部分を持たないかどうか判定される。これが当てはまると、ソース素材は、テンプレートのセクションに関連する部分を含まない。テンプレートのこれらのセクションに対応させることができるには不十分なソース間メタデータを有するソース素材の部分があることもある。本発明の他の具体例は、このような後の場合を、メタデータをこれらの部分にさらに対応させて扱うことを目的としている。

【0051】したがって、編集装置10が、ステップ91において占められていないテンプレートのセクションがある(YES)と判定すると、処理はステップ92に進み、編集装置10は、さらにメタデータを追加しようとする。上述したように、これには多くの方法がある。第1に、編集装置10は、ユーザにテンプレートのセクションにまだ対応していないソース素材の部分を表示し、ユーザにメタデータをこれらの部分に加えることを催促

する。ここで、認識処理装置36は編集装置10によって使用され、顔、声のような特徴を、ソース素材のこれらの部分において識別し、自動または手動によってメタデータをさらに生成することを容易にする。

【0052】一旦この処理が完了すると、ステップ93においてテンプレートは補給されたメタデータを使用するソース素材に再適用される。メタデータを補給することで、テンプレートを再適用するときに、以前に占められていなかったテンプレートの部分は、ここでこれらに対応するソース素材の部分の有することができる。

【0053】処理はステップ84に進み、編集コンテンツシーケンスは、ステップ93におけるテンプレートメタデータとソースメタデータとの比較に応じて、ソース素材の部分を配列することにより生成される。ステップ91で、占められていないテンプレートのセクションはない(N0)と判定されると、処理はステップ91からステップ84に直接に進む。

【0054】本発明の特定の具体例について説明してきたが、本発明はこれに限られるものではなく、本発明の趣旨を逸脱しない範囲で変更することができる。

【0055】

【発明の効果】上述したように、本発明によると、編集の前にソースコンテンツをレビューして編集する素材を理解する必要がなくなる。したがって、従来にはソースコンテンツをレビューするために費やしていた時間を削減することができ、編集作業の効率は著しく向上する。

【図面の簡単な説明】

【図1】図1は、編集装置の具体的な構成を示す図である。

【図2】図2は、本具体例のテンプレートを示す図である。

【図3】図3は、本具体例のテンプレートのソースコンテンツへの適用を示す図である。

【図4】図4は、本具体例のテンプレートのソースコンテンツへの適用を示す図である。

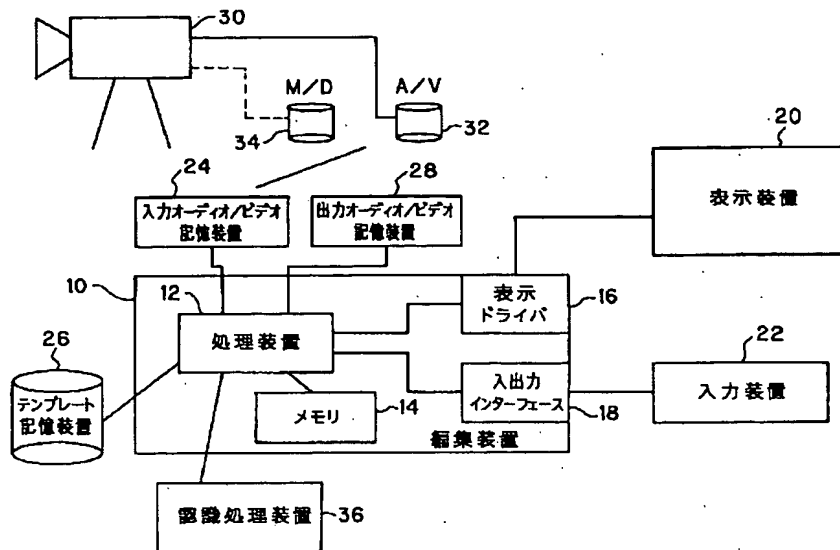
【図5】第1の具体例による編集装置の操作を示すフローチャートである。

【図6】第2の具体例による編集装置の操作を示すフローチャートである。

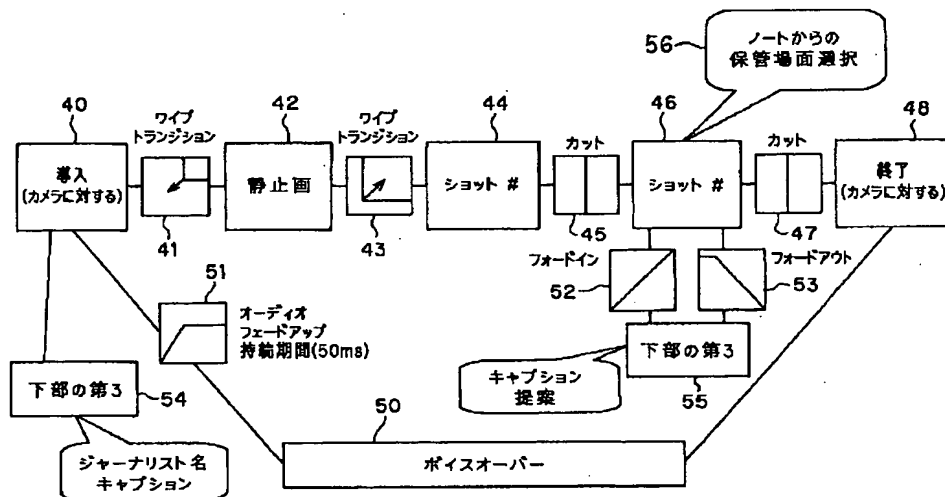
【符号の説明】

10 編集装置、12 処理装置、14 メモリ、16 表示ドライバ、18 入出力インターフェース、20 表示装置、22 入力装置、24 入力オーディオ/ビデオ記憶装置、26 テンプレート記憶装置、28 出力オーディオ/ビデオ記憶装置、36 認識処理装置

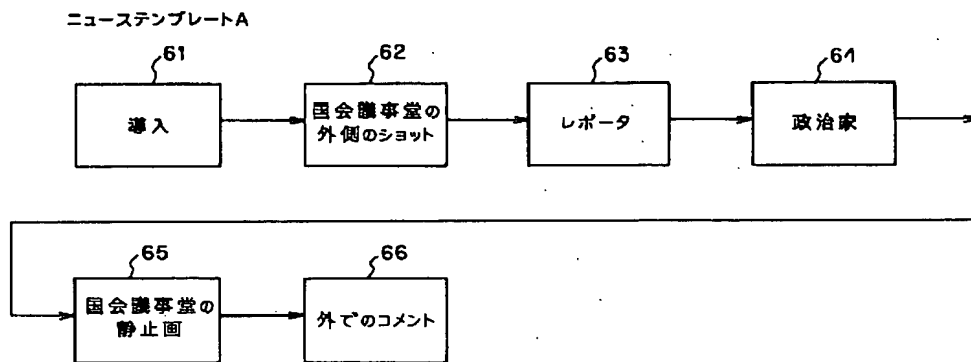
【図1】



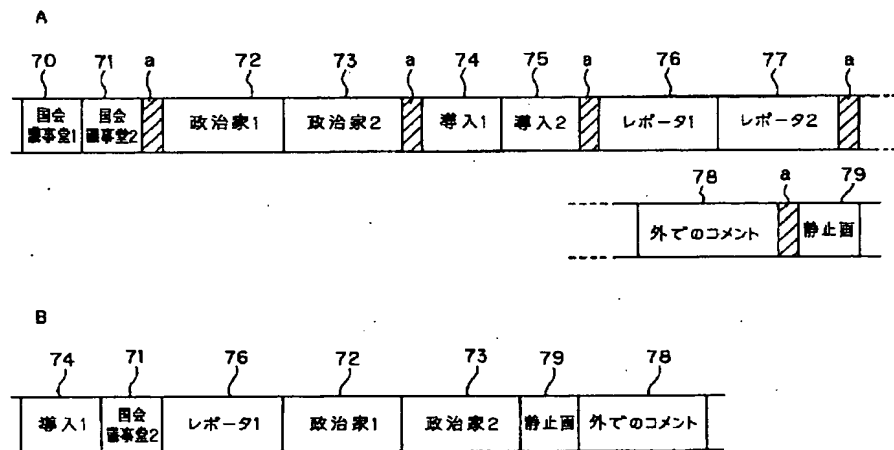
【図2】



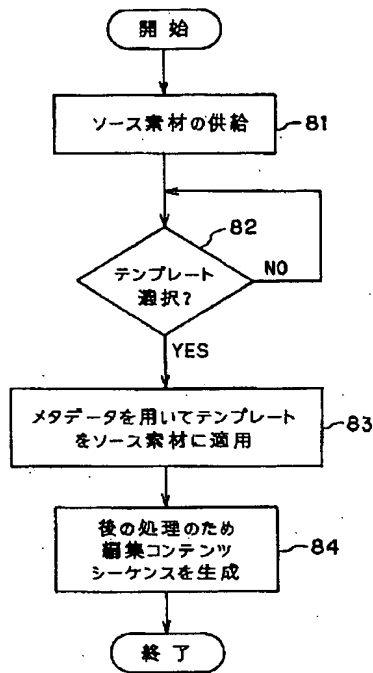
【図3】



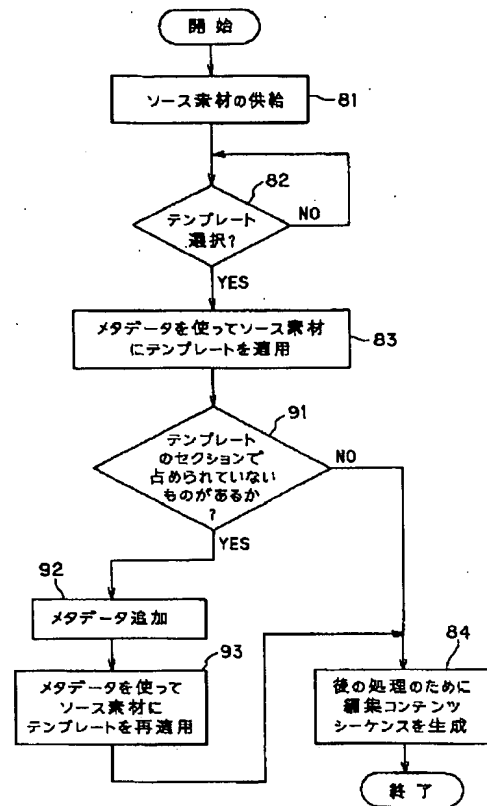
【図4】



【図5】



【図6】



フロントページの続き

(72)発明者 マクグラス、ジョン マーク  
イギリス国 ケーティー13 0エックスダ  
ブリュー サリー、ウエイブリッジ、ブル  
ックランズ、ザ ハイツ (番地なし) ソ  
ニー ユナイテッド キングダム リミテ  
ッド内

(72)発明者 コリンズ、アンドリュー  
イギリス国 ケーティー13 0エックスダ  
ブリュー サリー、ウエイブリッジ、ブル  
ックランズ、ザ ハイツ (番地なし) ソ  
ニー ユナイテッド キングダム リミテ  
ッド内

(72)発明者 ドリコット、レックス マーチン  
イギリス国 ケーティー13 0エックスダ  
ブリュー サリー、ウエイブリッジ、ブル  
ックランズ、ザ ハイツ (番地なし) ソ  
ニー ユナイテッド キングダム リミテ  
ッド内

PATENT ABSTRACTS OF JAPAN

(11)Publication number : 2001-143445

(43)Date of publication of application : 25.05.2001

---

(51)Int.Cl. G11B 27/031

H04N 5/91

---

(21)Application number : 2000-274132 (71)Applicant : SONY UNITED  
KINGDOM LTD

(22)Date of filing : 08.09.2000 (72)Inventor : MCGRATH JOHN MARK  
COLLINS ANDREW

DORRICOTT REX MARTIN

---

(30)Priority

Priority number : 1999 9921234

Priority date : 08.09.1999

Priority country : GB

---

(54) DEVICE AND METHOD FOR EDITING

(57)Abstract:

PROBLEM TO BE SOLVED: To provide an editing device which generates source metadata corresponding to an editing content sequence and the part of source contents by editing the source contents like video and/or audio raw materials.

SOLUTION: This editing device has a template storage device which stores a template prescribing a desirable style of contents, template metadata

prescribing the desirable style of the template, and a processor which applies the template to the source contents by comparing the source metadata with the template metadata and arrays the part of the source contents through the comparison to generate the editing contents sequence. Consequently, efficient technology for arraying the contents in the desirable style in a 1st cut is provided, so that an editor requires no time to understand the contents before editing the source contents. The user further advances alterations which seem to be proper from the 1st cut.

---

LEGAL STATUS [Date of request for examination]

[Date of sending the examiner's decision of rejection]

[Kind of final disposal of application other than the examiner's decision of rejection or application converted registration]

[Date of final disposal for application]

[Patent number]

[Date of registration]

[Number of appeal against examiner's decision of rejection]

[Date of requesting appeal against examiner's decision of rejection]

[Date of extinction of right]

\* NOTICES \*

JPO and INPIT are not responsible for any damages caused by the use of this translation.

1.This document has been translated by computer. So the translation may not reflect the original precisely.

2.\*\*\*\* shows the word which can not be translated.

3.In the drawings, any words are not translated.

---

CLAIMS

---

[Claim(s)]

[Claim 1] In the edit equipment which edits source contents and creates the source metadata corresponding to the parts of an edit contents sequence and the above-mentioned source contents The storage which stores the template metadata which specifies a desired style in the contents of a desired style, and a template, Edit equipment equipped with the processor which applies the above-mentioned template to the above-mentioned source contents, arranges the part of the above-mentioned source contents according to the



above-mentioned comparison, and creates an edit contents sequence by comparing the above-mentioned source metadata with the above-mentioned template metadata.

[Claim 2] Edit equipment according to claim 1 characterized by generating the edit contents sequence which arranges the part of the above-mentioned source contents and consists of selection, and setting in order the subset of the part of the above-mentioned source contents based on the comparison with the above-mentioned source metadata and the above-mentioned template metadata.

[Claim 3] The above-mentioned template is edit equipment according to claim 1 characterized by specifying the style of the above-mentioned request, and two or more discrete sections coordinated so that each discrete section which has corresponding template metadata might be specified.

[Claim 4] The above-mentioned source contents are edit equipment according to claim 3 characterized by consisting of the above-mentioned template metadata corresponding to the above-mentioned discrete section of a large number which specify the class of a video material and video material, and the discrete section which consists of metadata which specifies the class corresponding to a video material at least.

[Claim 5] The above-mentioned contents are edit equipment according to claim

3 characterized by consisting of the above-mentioned above-mentioned template metadata corresponding to the above-mentioned discrete section of a large number containing the metadata relevant to an audio material and the above-mentioned audio material.

[Claim 6] The edit equipment according to claim 1 carry out having the metadata processing element which one or more predetermined recognition tools access [ element ], and makes the above-mentioned source contents review with one or more [ above ] predetermined recognition tools in order to generate the source metadata corresponding to the above-mentioned source contents and the above-mentioned source metadata which is generated when applying the above-mentioned template to the above-mentioned source contents by comparing the above-mentioned template metadata for the above-mentioned source metadata, and is used by the above-mentioned processor as the description.

[Claim 7] The processing element of the above-mentioned metadata is edit equipment according to claim 6 characterized by enabling a dialogue with a user between generation of the above-mentioned source metadata according to a user input.

[Claim 8] The above-mentioned template is edit equipment according to claim 1 characterized by specifying the template metadata which specifies the style of

the above-mentioned request in contents and the above-mentioned template.

[Claim 9] Edit equipment according to claim 8 characterized by consisting of two or more discrete sections coordinated in order to specify the style of the above-mentioned request, and each discrete section which has corresponding template metadata.

[Claim 10] The step which receives the source metadata corresponding to the parts of the source contents by which are the edit approaches which operate edit equipment, edit source contents and create edit contents, and (a) edit is carried out, and the above-mentioned source contents, (b) The step which offers the template which specifies the template metadata which specifies a desired style in contents and the above-mentioned template, (c) The step which applies the above-mentioned template to the above-mentioned source contents by comparing the above-mentioned source metadata with the above-mentioned template metadata, (d) The edit approach of having the step which arranges the part of the above-mentioned source contents according to a comparison at a step (c), and generates the above-mentioned edit contents sequence.

[Claim 11] It is the computer program which operates the edit equipment into which source contents are edited into and the source metadata corresponding to the parts of an edit contents sequence and the above-mentioned source contents is edited. The step with which the above-mentioned computer program

searches the template metadata which specifies a desired style to the above-mentioned edit equipment in contents and a template from (a) storage, (b) The step which applies the above-mentioned template to the above-mentioned source contents by comparing the above-mentioned source metadata with the above-mentioned template metadata, (c) Computer program characterized by performing actuation which consists of a step which arranges the part of source contents according to a comparison at a step (b), and creates the above-mentioned edit contents sequence.

[Claim 12] It consists of two or more discrete sections coordinated so that a desired style may be specified in the template used with the edit equipment into which source contents are edited into and the source metadata corresponding to the parts of an edit contents sequence and the above-mentioned source contents is edited. The above-mentioned discrete section is a template characterized by having corresponding template metadata.

---

## DETAILED DESCRIPTION

---

[Detailed Description of the Invention]

[0001]

[Field of the Invention] This invention relates to the edit equipment which edits

source contents (source content) and creates an edit contents sequence (edited content sequence) and an approach, and a computer program list at a template. Source contents are video and/or an audio material (video/audio material).

[0002]

[Description of the Prior Art] From the former, a burden is a large activity and edit of source contents is performed by the skilled editor with edit equipment (editorial suite). Edit equipment (editing system) is used for usually editing source contents under management of an editor. The present edit equipment has the refined tool and an editor can process source contents by the complicated edit effectiveness.

[0003]

[Problem(s) to be Solved by the Invention] However, the editor has required considerable time amount, in order to understand the material into which source contents are reviewed and (review) edited before such edit. The effectiveness of an editing task is spending time amount, in order to understand source contents first, and it is getting worse.

[0004] This invention is proposed in view of the above-mentioned actual condition, and aims at providing with a template edit equipment which raises edit effectiveness and an approach, and a computer program list by reducing the time amount spent in order that an editor may understand a material before edit.

[0005]

[Means for Solving the Problem] This invention offers the edit equipment which edits source contents and creates the source metadata (source metadata) corresponding to the parts (portion) of an edit contents sequence and source contents. This edit equipment is equipped with the store (storage) which stores the template metadata (template metadata) in the template (template) which specifies the contents of a desired style, and a desired style, and the processor (processing unit) which applies a template to source contents by comparing source metadata (source metadata) with template metadata, arranges the part of source contents according to this comparison, and creates an edit contents sequence.

[0006] This invention offers the template metadata which specifies a desired style in the template and template which specify the style of a request of contents. In this example, a template consists of two or more discrete sections coordinated so that a desired style may be specified (link), and each discrete section has corresponding template metadata. Therefore, according to this viewpoint of this invention, many templates are prepared, a specific format required for an edit contents sequence is specified, and these templates are chosen so that it may be suitable for use of the edit equipment of this invention.

[0007] This invention offers the edit approach of operating edit equipment so that

source contents may be edited and edit source contents may be created. This edit approach (a) The step which receives the source metadata corresponding to the parts of the source contents edited and source contents, (b) The step which offers the template metadata which specifies a desired style in contents and a template, (c) It has the step which applies a template to source contents by comparing source metadata with template metadata, and the step which arranges the part of source contents according to the comparison of the (d) step (c), and creates an edit contents sequence.

[0008] This invention offers the computer program which operates edit equipment so that the source contents metadata corresponding to a part may be edited into source contents and source contents and an edit contents sequence may be created. The step with which a computer program searches the template metadata which specifies the style of the request in the template and template which specify a desired style to edit equipment from (a) storage, (b) The step which applies a template to source contents by comparing source metadata with template metadata, (c) The step which arranges the part of source contents according to the comparison in a step (b), and creates an edit contents sequence is performed.

[0009]

[Embodiment of the Invention] Hereafter, the gestalt of operation of a template in

the edit equipment concerning this invention and an approach, and a computer program list is explained to a detail, referring to a drawing.

[0010] Drawing 1 is the block diagram showing the concrete configuration of the edit equipment 10 which applied this invention. As shown in this drawing 1, edit equipment 10 is equipped with the processor 12 which receives source contents and performs editing operation to source contents (processing unit), for example, a microprocessor, in order to create edit contents. Internal memories 14 are a read only memory (read only memory: ROM), random access memory (random access memory: RAM), etc., and store the software performed with a processor 12 with the data used with a processor 12.

[0011] Edit equipment 10 is equipped with the display driver 16, and the display driver 16 outputs a status signal to the display 20 connected to edit equipment 10. Moreover, edit equipment 10 is equipped with an input/output interface 18, and an input/output interface 18 manages the interface which connects an input device 22, for example, a keyboard, a mouse, a jog shuttle dial, a pen, a tablet, etc. to edit equipment 10. The user of edit equipment 10 can have a dialog with the edit processing performed by edit equipment 10 through a display 20 and an input unit 22 (interact).

[0012] By this example, it is supposed that they are the source contents edited by edit equipment 10 the audio / video (audio/vide: henceforth A/V) material



which was outputted from the digital camera 30. Although, as for an A/V material, analog data is also created, since the edit equipment 10 of this example operates with digital data, an analog A/V material is changed into digital data before [ edit equipment 10 ] being inputted.

[0013] It is stored in the suitable medium 32, for example, a digital tape, or a digital disk when an A/V material is generated by the digital camera 30. When an A/V material is created first, a certain metadata is made to correspond in the latest digital technique by the A/V material. By this example, such metadata is stored in a digital tape or a disk as a suitable medium 32. In addition, although metadata is recorded on the same physical memory medium 32 as a corresponding A/V material, it is also recordable also on another physical record medium 34.

[0014] Additional metadata can also be added by adding metadata at a certain time before an A/V material is sent to edit equipment 10 manually for example. When an A/V material is edited, an A/V material is stored in the input A/V store (input A/V store) 24, for example, a digital tape player, a digital disk player, or a server with corresponding metadata (below, it is called source metadata.).

[0015] By this example, one or more templates are offered from a store, for example, the store accessed through a database or a network. Here, although expressed with drawing 1 as a component with the separate template storage 26,

a template is also storable in memory 55 locally in edit equipment 10. It is admitted that many kinds of edit materials, for example, a news report, follow the similar pattern. Therefore, by this example, each template specifies the style of a request of an A/V material, and template metadata specifies a desired style in each template by it.

[0016] Edit equipment 10 searches a suitable template from the template store 26, and is made as [ apply / by comparing template metadata with the source metadata corresponding to a source A/V material / to the source A/V material supplied from the input A/V store 24 / a template ]. This activity is done by the processor 12, and according to the comparison of source metadata and template metadata, a source A/V material is arranged so that an edit contents sequence may be created. An edit contents sequence is outputted to the output A/V storage 28 from a processor 12. This output can be further changed by the user of edit equipment 10. The output A/V store 28 as well as the input A/V store 24 consists of a digital tape, a digital disk player, or a server.

[0017] In this example, it is assumed that a template is chosen by the user of edit equipment based on an understanding of the user to the A/V material to edit. For example, a user can give directions explicit about selection of a template from those who are asking for edit of an A/V material. Following [ for example, ], the user of edit equipment 10 receives a certain A/V material related to a news

matter with directions of using the template for the specific news stored in the template storage 26. In addition, a user can also decide the template to be used by himself based on the type of the A/V material to edit. Although a template is specifically manually chosen from the template storage 26, it is also considered like automatic selection of the template by edit equipment 10 that the metadata corresponding to the A/V material to edit directs to a processor 12 which template is used.

[0018] By this example, edit equipment 10 is connectable with many recognition processors 36 which have the gestalt of one or more plug-in recognition tools. These recognition tools are used for assistance of generation of metadata by edit equipment 10. For example, the recognition tool which detects a face is used in order to look for the face which scans source A/V and is contained in video. Once these faces are identified, the user of edit equipment 10 will be shown these and they will urge him to input the metadata related to discernment of the face by which the user was detected. In addition, if a database with any of a recognition tool or edit equipment 10 suitable for people's face they are can be accessed, automatic recognition of a face can also be performed. Similarly, voice is recognized and the recognition tool which can perform the addition of the metadata relevant to the audio material in the A/V source is also offered.

[0019] Next, the template used by edit equipment 10 is explained using drawing

2 . Drawing 2 is drawing showing the concrete template stored in the template storage 26. The template which describes an A/V sequence consists of 5 main sections 40, 42, 44, 46, and 48. The 2nd section 42 consists of one or more still pictures (still), the 3rd section 44 gives the scene of a specific shot (shot), the 1st section 40 consists of a material which offers installation (introduction), and the last section 48 offers [ the 4th section 46 gives the scene of a shot further, and ] a termination sequence. Furthermore, the voice-over (voice over) of a section 50 is used in order to give a material an audio voice (audio voice) and to give the audio corresponding to sections 40, 44, and 46.

[0020] (Video wipe (video wipe), a page turn (page turn), the audio fade (audio fade), cross fade (cross fade), etc. are needed) Much specific edit effectiveness, keying, and/or technical adjustment correspond to each section, and a user is reminded of an input during generation of an edit contents sequence, using which of the effectiveness and the keying which these-correspond, or adjustment. For example, a section 40 corresponds to the edit effectiveness which shows the box wipe (box wipe) to the lower left from the upper right applied to the last of installation (it is shown in block 41 like), and the section 42 supports the edit effectiveness which shows the box wipe to the upper right from the lower left applied to the last of a still picture (it is shown in block 43 like). Similarly, a straight cut (straight cut) is specified among sections 44 and 46 and

among sections 46 and 48, as shown among blocks 45 and 47. Moreover, as shown in block 52, an audio fade rise (audio fade up) is specified as the beginning of a section 50, and as shown in blocks 52 and 53, video fade-in (video fade in) and video fade-out (video fade out) support the caption of a section 46. For example, a caption is proposed by the user as shown in the blocks 54 and 55 corresponding to sections 40 and 46, respectively.

[0021] Each section has the matter of one or more corresponding metadata. Such metadata is called template metadata below. The template metadata corresponding to a specific section is used for specifying the class of contents corresponding to the section. There are many classes corresponding to each section of a template of metadata. The next table 1 shows an example with the class of metadata.

[0022]

[Table 1]

ビジネス情報：カメラマン、ジャーナリスト、割当番号、版權	トランジションマーク：カット、ディゾルブ、フェードの自動識別及びカテゴリ化
装備仕様：カメラシリアル番号、レンズパラメータ、技術特徴	ショット構成マーク：ズーム及び広角度マーク
素材識別：固有素材識別子（UMID）	雑音マーク：高周雑音（おそらく高電力低周雑音）又は低周雑音（すなわち無感部分）
位置マーク：GPS、時間帯、固定位置（例えば、イギリス国会の議場のマイクロフォン、ゴール端のカメラ）	テンポ／ビートマーク：リズム、周期的音響イベント
ユーザ定義マーク：グッドショットマーク、制作コメント	色／コントラスト変化：カメラの閃光、電光効果によって引き起こされる。
外部運動：ズーム又はパニングによる大域的な運動	ピッチ（音の高さ）／ボリューム（音の太さ）変調：これは、音楽や対話のような長期間のオーディオイベントに当てはまる。
内部運動：スポーツにおける、又は人の動き（すなわち、「比較的」遅く変化する運動を伴う人のような凝集した被写体）のような被写体の軌跡、	オーディオパターン／不連続：リズム、テンポ、歩調のようなもの。

[0023] Recently, it is interested in generation of metadata and the metadata which describes contents. It is thought that the correspondence relation between meter data and contents is useful in the case of storage (archive). Metadata is used for the assistance in processing of the retrieval (search) after desired contents, and retrieval (retrieve). Much specification is developed by metadata like SMPTE-EBU (Society of Motion Picture and Television Engineers-European Broadcast Union) and MPEG-7 (Motion Pictures Expert Group). Here, Motion Pictures Expert Group means ISO/IEC standards body SG29\*\*WG11.

[0024] the metadata of a constant rate -- for example, (it is shown that it is a good shot) -- good, a shot marker (good shot marker), a sound recording (it is shown where record was started / stopped) marker, a GPS (global positioning

system) location, time, a proper material identifier (unique material identifier: henceforth UMID), a camera setup, a focus, a zoom, etc. can be automatically added to the source. In addition, an advance of the technique which carries out automatic extracting of the metadata from contents is large, for example, by using recognition tools, such as a face and a feature detection tool, and a speech recognition tool, etc., the description of contents can be identified and suitable metadata can be added.

[0025] This invention uses the latest development of the metadata (henceforth source metadata) technique corresponding to source contents, and is aiming at improving the effectiveness of edit processing. Concretely, by this invention, a template specifies the style of a request of the template metadata which specifies a desired style in contents and a template. Template metadata essentially expresses selection of source metadata, or the adaptation criteria of source metadata. as [ apply / the template stored in a hard disk, a floppy (trademark) disk, CDROM, etc. is searched by edit equipment, and / by comparing source metadata with template metadata / to source contents / a template ] -- it is carried out. The part of source contents is arranged according to this comparison, and an edit contents sequence is created.

[0026] In this invention, many kinds of edit materials, for example, a news report, a sport clip / outline, a video montage, documentary, etc. shall be followed by the

quite similar pattern. For example, the report of the discussion in Parliament starts in installation (introduction), progresses to the shot (shot) of the outside of the Diet Building, the cut (cut) of discussion of a politician, and a cut of everything but the Diet Building again, and returns to the reporter who finally describes the comment of an end. The edit effectiveness applied to a shot or the edit effectiveness applied between shots also follows the similar pattern. This invention uses that such a fact and metadata are easy to come to hand, and offers the edit equipment with which the template metadata in these templates specifies a desired style about the template created to the style of various requests. A desired style specifies the descriptions, such as a general sequence of a shot, criteria used for choosing a part from source contents, and the edit effectiveness applied.

[0027] In order to create an edit contents sequence, the selected template is applied to source contents by comparing source metadata with template metadata. In order to understand contents before a user edits source contents, it becomes unnecessary for this edit contents sequence to spare time amount, since source contents provide an editor with "the 1st cut (first cut)" arranged in the desired style. It can progress to modification considered for a user to be suitable from this 1st cut.

[0028] In order to arrange source contents and to create an edit contents



sequence, one or more steps are required. For example, the array of source contents contains steps, such as record of the part of source contents, selection of the part from source contents, application of the edit effectiveness to source contents, and insertion of a caption, one or more. By this example, the creation of an edit contents sequence which arranges the part of source contents has the selection or sequencing of the subset of the part of source contents based on the comparison with source metadata and template metadata at least. It follows and for example, this example shows whether which Taegu of source metadata is desirable using the good shot marker corresponding to Taegu including multiplex "Taegu (take) of the shot with the same source contents." Good, in creation of an edit contents sequence, if a shot marker is specified by template metadata, the comparison with template metadata and source metadata will be used in order to choose only Taegu of the request used from a specific shot. By this approach, it is thought that an edit contents sequence quite shorter than original source contents (original source content) is created.

[0029] Although a template takes various forms, by this example, two or more discrete sections (discrete section) which cooperate in order that a template may specify a desired style are specified, and each discrete section has corresponding template metadata. Each discrete section is considered to specify the type of the contents of the request corresponding to the part of a desired

sequence by this example, and one or more items (item) of template metadata specify the class of contents corresponding to this discrete section corresponding to each discrete section.

[0030] In this example, source contents consist of a video material, many discrete sections specify the class of video material, and the template metadata corresponding to such a discrete section consists of metadata which specifies the class of corresponding video material at least. For example, one discrete section is made into "quiescence (still)" shot, and the metadata to this discrete section is a class which specifies the quiescence shot for example, in source contents. A quiescence shot [ in / as compared with source metadata / in the template metadata of this discrete section / source contents ] is specified by this approach, and one or more parts of the source contents showing these quiescence shots are made to correspond with it by the discrete section used for creation of an edit contents sequence.

[0031] For example, in addition to a video material, apart from a video material, contents consist of an audio material and the template metadata corresponding to many discrete sections contains the metadata relevant to an audio material. Such metadata expresses a high background noise (background noise), a low background noise, a rhythm, a periodic sound event, etc., for example, and when the part of the source contents which contain the audio of a desired type

for template metadata as compared with source metadata is specified, they can be used for it.

[0032] As mentioned above, metadata is manually added in a certain phase after being added to source contents in the source or recording source contents. In addition, in order for edit equipment to generate source metadata and to make it correspond to source contents by this example, one or more predetermined recognition tools are accessed, it is generated in this way with the metadata processing element which makes source contents review with one or more recognition tools, and when applying a template to source contents by comparing source metadata with template metadata, it consists of source metadata used by the processor. A metadata processing element (metadata processing element) is built into hardware, firmware, and/or software so that it may be suitable for access to one or more predetermined recognition tools (recognition tool), and control of generation of the metadata which used these tools. The recognition tool which follows, for example, identifies a face is used so that the face which scans video and is contained in video may be looked for. Once these faces are identified, recognition of the face by automatic or hand control will be used, a related person will be specified, and the suitable metadata according to recognition will be added.

[0033] In this example, a metadata processing element can carry out a dialogue

(interaction) with a user between generation of source metadata according to a user input. Therefore, in the above-mentioned example, a user is shown the found-out face and a user is urged to identify the person shown in the metadata about these images, for example, an image.

[0034] If a specific template is chosen, the processor 12 of edit equipment 10 will apply a template to a source A/V material, in order to generate an edit contents sequence. This is performed by comparing source metadata with template metadata, in order to make each part of a source A/V material correspond to the specific section of a template. For example, the specific part of a source A/V material has the source metadata corresponding to the template metadata corresponding to a section 42, and specifies that these parts of an A/V material are related to a still picture. Similarly, other parts of an A/V material have the source metadata corresponding to the template metadata corresponding to the section 44 which specifies the 1st specific shot. Moreover, it has the source metadata which specifies that the part with a source material relates to a voice-over material, and has the source metadata relevant to a section 50 which carries out thing specification, and correspondence with the template metadata relevant to the section 50 of metadata is used in order to make a source material correspond to the section 50 of a template.

[0035] According to the comparison with template metadata and source

metadata, the part to which a source A/V material relates is arranged according to a comparison, and generates the edit contents sequence according to the general style specified by the template. It is made for this to follow the style which set these parts in order again including selection of the subset of the part of a source material preferably, and was specified by the template. What kind of specified edit effectiveness is applied at this time. Therefore, by choosing a suitable template, the user of edit equipment is automatically provided with the 1st cut of an edit sequence, the part of a request of source contents is chosen and a sorting eclipse and the edit effectiveness are applied. Therefore, a user can judge whether it is necessary to review an edit contents sequence and to change it further.

[0036] Since the scene (footage) of the suitable addition for use is searched, a certain tag can also be made to correspond to a specific section so that it may look for a storage material (archive material) with edit equipment 10. This retrieval is started by comparing the metadata corresponding to the specific template metadata and the storage metadata of a section to which a tag corresponds in order to search the suitable scene relevant to this section. If the example of drawing 2 is referred to, this processing is shown by the block 46 corresponding to a section 46, and it is shown that the source metadata corresponding to the source A/V material of the form of notes has selection of a

specific storage scene.

[0037] Application to the source contents of a template is explained using the example of drawing 4 from drawing 3 . Drawing 3 shows the news template A specified by the maker of the specific source A/V material which is related to a news matter. A desired style is specified by six sections 61, 62, 63, 64, 65, and 66 according to the news template A. In relation to the installation about the news item which reports the 1st section 61, as for the 5th section 65, the 2nd section 62 gives [ give / as for the 3rd section 63 / by giving the shot outside the Diet Building / Leh Porter's shot ] the still picture of the Diet Building by giving a politician's shot, as for the 4th section 64, and the last section 350 gives the comment of the end outside a building about an old topic. Like the template described with reference to drawing 2 , although the specific edit effectiveness etc. corresponds to each section, since it is easy, these are omitted by drawing 3 . Each section determines whether it has the item of the template metadata of corresponding a large number, source contents are reviewed, and the specific part of source contents is made to correspond by the section of which template.

[0038] A of drawing 4 shows the original source A/V material so that it may appear in a source medium. A source material consists of each shot or ten parts relevant to Taegu by which the capture was carried out in this example. There is a miscellaneous material which is concretely uninterested like the material

recorded between camera setup between a certain parts, and this material is expressed by the field a of drawing 4 . It sets to the edit equipment of a specific prior art, and which part is used for the specific part of the last version of an edit contents sequence, and in order to decide, the user (followed on someone who was often concerned with work of source contents, for example, a journalist) of edit equipment has to scan the original source several times, and has to understand all the contents of a source material. However, by this example, someone who was concerned with work of source contents, for example, a producer, specifies that an edit contents sequence follows the news template A. The news template A is chosen from the template storage 26 used by edit equipment 10 according to this.

[0039] Source metadata is made to correspond a template by various parts of source contents before [ source contents ] applying. When the capture of the source metadata is carried out first, it is made to correspond by contents, or it is added by manual insertion of suitable metadata later. With corresponding source metadata, if edit equipment 10 receives source contents, edit equipment 10 will be comparing the template metadata corresponding to each section of Template A with the source material corresponding to the part of a source material, and will apply the news template A to a source material.

[0040] Thus, in drawing 3 , as the introductory section 61 is shown in drawing 4

in the review of source metadata including the template metadata of a type which generally specifies a part for the induction in a source material, it is found out that the item of the template metadata corresponding to a section 61 corresponds to the source metadata corresponding to the parts 74 and 75 showing two Taegu of installation. For example, parts 74 and 75 are introducing by insertion by the handicraft of the source metadata which specifies these parts as installation, and are identified as it is Taegu.

[0041] In addition, for example, these parts contain which [ of installation / two / of Taegu ] are desirable, and the metadata to identify by for example, a good shot marker's use. If the template metadata corresponding to a section 61 can also contain the template which supports discernment of the part marked as good in the source material, Taegu of a request of installation will be chosen by applying a template to source contents. Therefore, in the example of drawing 4 , such source metadata is added and it is supposed that the part 74 showing the 1st desirable Taegu of installation 1 is desirable. Similarly, the part 71 supports the section 61.

[0042] Similarly, the template data corresponding to a section 62 are specifically metadata, and support specification of the shot of the Diet Building. For example, the metadata which identifies the GPS location of the Diet Building supports the section 62. When reviewing source contents and it is found out that GPS



metadata supports parts 70 and 71, these parts are the purposes which determine the sequence of the part in the edit contents sequence made, and must be made to correspond by the part 62. In addition, in this example, source metadata specifies one of the shots noting that it is more good, and the shot expressed by the part 71 in this case and template metadata include the class which identifies a good shot through source metadata. Therefore, a shot 71 corresponds to a section 62.

[0043] Similarly, metadata is added to the source contents which specify the part of the source contents related to a reporter's Taegu, and the part of the contents related to a politician's Taegu about the sections 63 and 64 related to a reporter and a politician, respectively. Although this is applied directly, it is used by the user in order to identify these parts of the source material which offers the shot of a face, and he is provided with a recognition tool like the face detection tool which carries out label attachment of these parts by suitable metadata, and specifies the part related to a reporter, and the part related to a politician. Therefore, when reviewing a source material, edit equipment 10 specifies that parts 76 and 77 correspond to a section 63 appropriately, and parts 72 and 73 correspond to a section 64 appropriately. However, in this example, since a good shot marker corresponds to a part 76, priority is given over a part 77 and he is chosen from it.

[0044] In this example, a good shot marker does not correspond to parts 72 and 73, in a desirable example, both locations correspond to a section 64 and a user can do two selections of Taegu later. This is applied when it does not include the type of the template metadata which can perform discernment of a shot with a desirable section 64, or Taegu like [ when not including the class in which for example, template metadata supports a good shot marker's etc. specification ].

[0045] Suitable metadata corresponds to sections 65 and 66, and supports discernment of every part of a static image or the source contents related to each of a comment [ in outside ]. Therefore, the edit source contents as which it is applied to the source contents the news template A shown in drawing 3 was indicated to be to A of drawing 4 , and the source contents by which the rearrangement was carried out are specified by B of drawing 4 are obtained by comparing template metadata with source metadata.

[0046] The edit contents sequence automatically generated by edit equipment 10 is displayed that the user of edit equipment can perform required adjustment further by comparing source contents with the news template A. In this example, priority is given at Taegu/shot of an alternative with a desirable part with source contents, and since the rearrangement of the part of various source contents dropped and fragmented is carried out to the suitable sequence of the last edit material, in order to understand contents before a user edits source contents, it

is considered that it becomes unnecessary to spare time amount. From this 1st cut, a user can progress to the creation of the further modification considered to be suitable further. The effectiveness of edit processing is remarkably improved by this.

[0047] Drawing 5 is a flow chart which shows the processing performed with edit equipment 10 in the 1st example. In step 81, a source material is supplied, for example, it is read into the input A/V storage 24. In step 82, edit equipment 10 judges whether the template is chosen or not. By this example, when a user makes suitable selection through an input unit 22, it is carried out manually. In addition, as mentioned above, since the metadata which specifies the template to be used and can be directly recognized with edit equipment 10 has already supported the source material, edit equipment 10 can choose a desired template automatically.

[0048] Processing progresses to step 83, and edit equipment 10 searches a template from the template store 26, and applies it to a source material by comparing template metadata with source metadata. In step 84, an edit contents sequence is created for the review by the user as a result of a comparison. As mentioned above, an edit contents sequence is created by arranging the part of a original source material according to the comparison of the template metadata performed at step 83, and source metadata. For example, although this

processing generates an edit contents sequence including selection and sequencing of the part of a source material, the application of the edit effectiveness etc. includes it.

[0049] In drawing 5 , the source material is considering as the thing to the source material of the template containing suitable metadata to apply. Drawing 6 shows the 2nd example to which source metadata is supplied between edit processings.

[0050] The first three steps 81, 82, and 83 of drawing 6 are the same as that of what was shown by drawing 5 above. However, once a template is applied to a source material by comparing template metadata with source metadata, it will be judged whether what the part of a template is not occupied for in step 91 (not populated), i.e., no section, has the part of the source material corresponding to these. If this is applied, a source material does not contain the part relevant to the section of a template. There may be a part of the source material which has the metadata between the sources inadequate for the ability to make it correspond to these sections of a template. Other examples of this invention aim at making metadata correspond to these parts further, and treating a such next case.

[0051] Therefore, if edit equipment 10 judges with there being a section of a template which is not occupied in step 91 (YES), processing tends to progress to step 92 and, as for edit equipment 10, it tends to add metadata further. As

mentioned above, there are many approaches in this. Edit equipment 10 displays the part of the source material which does not correspond yet into the section of a template on a user, and reminds the 1st of adding metadata to these parts at a user. Here, it is made easy that the recognition processor 36 is used by edit equipment 10, identifies a face and the description like voice in these parts of a source material, and generates metadata further with automatic or hand control.

[0052] Once this processing is completed, in step 93, a template will be re-applied to the source material which uses the supplied metadata. When re-applying a template by supplying metadata, the part of the template which was not occupied before can have the part of the source material corresponding to these here.

[0053] Processing progresses to step 84 and an edit contents sequence is generated by arranging the part of a source material according to the comparison with the template metadata and source metadata in step 93. At step 91, if the section of a template which is not occupied is judged to be (NO) which is not, processing will progress to step 84 directly from step 91.

[0054] Although the specific example of this invention has been explained, this invention is not restricted to this and can be changed in the range which does not deviate from the meaning of this invention.

[0055]

[Effect of the Invention] It becomes unnecessary to understand the material into which source contents are reviewed and edited before edit according to this invention, as mentioned above. Therefore, the time amount spent in order to review source contents to the former can be reduced, and the effectiveness of an editing task improves remarkably.

---

## DESCRIPTION OF DRAWINGS

---

[Brief Description of the Drawings]

[Drawing 1] Drawing 1 is drawing showing concrete \*\*\*\*\* of edit equipment.

[Drawing 2] Drawing 2 is drawing showing the template of this example.

[Drawing 3] Drawing 3 is drawing showing application to the source contents of the template of this example.

[Drawing 4] Drawing 4 is drawing showing application to the source contents of the template of this example.

[Drawing 5] It is the flow chart which shows actuation of the edit equipment by the 1st example.

[Drawing 6] It is the flow chart which shows actuation of the edit equipment by the 2nd example.

[Description of Notations]

10 Edit Equipment, 12 Processor, 14 Memory, 16 Display Driver, 18 Input/output  
Interfaces, 20 Indicating Equipment, 22 Input Device, 24 Input Audio / Video  
Store, 26 Template Store, 28 Output Audio / Video Store, 36 Recognition  
Processor